

questioning, automatic speech recognition, and text-independent speaker recognition techniques to be utilized to *verify the identity of a user*. Specifically, Kanevsky teaches a security system that, through an iterative process, compares user responses to a user database, referred to as speaker candidates, of non-acoustic information and/or an acoustic user model to perform the verification/identification of the user requesting access to a service/facility (Kanevsky, col. 5, lines 34-46). The system first performs an automatic enrollment process by obtaining name, address and whatever other identification is required, this information is referred to as indicia, for building the user model and database used for future identification and verification of the user (Kanevsky, col. 8, lines 23-22). The identification process includes receiving a spoken utterance containing speaker indicia, decoding the spoken utterance to produce a decoded utterance indicia, accessing a database corresponding to a determined speaker candidate of the decoded utterance indicia based on similar indicia within the database, querying the speaker for an additional spoken utterance to obtain additional indicia, receiving and decoding the additional utterance, verifying accuracy of additional decoded utterance indicia against the accessed database, taking a voice sample from the utterances of the speaker and processing the voice sample against an acoustic voice model attributable to the speaker candidate, generating a score corresponding to the accuracy of the decoded answers and closeness of the match between the voice sample and the model, and comparing the score to a predetermined threshold value (Kanevsky, col. 3, lines 22-45). A voice classification module and a text-independent **speaker recognition module** matches the voice sample of the speaker to voice prints of specific words, the indicia, stored within the accessed acoustic model of the speaker candidate (Kanevsky, col. 6, line 67 to col. 7, line 5; col. 9, lines 46-58; col. 10, lines 56-60). Specific words within the voice sample are determined by a speech recognition device (Kanevsky, col. 6, lines 4-11). Summarily, Kanevsky teaches a system that improves a speaker identification process. More particularly, Kanevsky teaches a system that determines indicia words from a prompted voice sample using a speech recognition device and compares these indicia words to voice prints of corresponding indicia words attributed to a previously identified speaker stored within an acoustic model. *The comparison is made in order to verify the identity of a speaker of the voice sample. Kanevsky does not teach a system to modify a speech recognition system for the purpose of improving the recognition of indicia words within the voice sample.*

In contrast to the teachings of Kanevsky, the speech recognition system of the present

invention provides for speaker-specific acoustic models to be used in the **speech recognition process**. Multiple users can access the same application, each user having an individualized speaker-specific acoustic model which is stored, retrieved from storage, and modified according to samples of the specific speaker's speech. By utilizing speaker-specific models which are uniquely tailored to the individual user, the speech recognition system of the present invention greatly improves the accuracy of speech recognition over that of a generalized speech recognition system. Further, the speaker-specific acoustic model does not require direct user feedback to be modified. The present invention eliminates the inconvenience of requiring user feedback and as a result improves efficiency by automatically modifying the speaker-specific acoustic models based on the received samples of the specific speaker's speech. Kanevsky teaches a system to modify a **speaker identification process**. Kanevsky does not teach a system to modify a **speech recognition system**. Further, Kanevsky does not teach a system to modify and subsequently store a speaker-specific speech recognition system.

Within the Office Action, it is stated that Kanevsky does teach modifying a speech recognition system according to a sample of the speaker's speech thereby forming a speaker-specific modified speech recognition system, storing a representation of the speaker-specific modified speech recognition system in association with the identification of the speaker, and storing a representation of the speaker-specific modified speech recognition system to recognize speech during a subsequent remote session with the speaker. To support this assertion, column 8, lines 45-55 of Kanevsky is cited. The Applicant respectfully disagrees with this reading of those portions of Kanevsky. In lines 44-55 of column 8, Kanevsky teaches that, in the case where no previous user acoustic model exists for the caller, the system collects voice samples from the caller's answers to a plurality of questions and builds a new user acoustic model for the caller therefrom. Then, the next time the caller calls, the server 22 asks a few random questions from the database which are then processed through an automatic speech recognition (ASR) 28. In addition, the server 22 uses the text-independent speaker recognition module 52 to match identified indicia from the ASR 28 to already stored indicia in the acoustic model to **verify the caller's identity**. The speaker recognition module 52 is separate and distinct from the user acoustic model. As such, the collected voice samples are used to build the user acoustic model, not to modify the speaker recognition module 52. Further, the collected voice samples are not applied to the ASR 28 in any manner. Instead, the ASR 28 is used to recognize appropriate

indicia from the responses to the few random question which were posed to the caller when the caller calls subsequent to generating the new acoustic model. Then, the recognized indicia are compared to stored indicia corresponding to the voice samples within the user acoustic model for the purpose of verifying the caller's identity by the speaker recognition module 52. **At no point in Kanevsky are the collected voice samples used to modify the ASR 28.** More specifically, Kanevsky teaches one module for recognizing indicia from within a voice sample, the ASR 28 (Kanevsky, col. 6, lines 7-11), and another separate module for comparing the recognized indicia words to stored voice prints within the user acoustic model for verifying the identity of the speaker of the voice sample, the speaker recognition module 52 (Kanevsky, col. 10, lines 56-60). The stored voice prints correspond to the voice samples obtained when the new user acoustic model was generated. In fact, Kanevsky teaches that the ASR 28 used to perform the necessary speech recognition of the indicia words is a conventional speech recognition device (Kanevsky, col. 13, lines 48-61). It is this type of conventional speech recognition technology that the present invention improves upon (Specification, page 2 lines 3-4; page 2, lines 8-16; and page 2 lines 30-31).

The independent Claim 1 is directed to a method of adapting a speech recognition system. The method of Claim 1 includes the steps of obtaining an identification of a speaker, obtaining a sample of a speaker's speech during a first remote session, recognizing the speaker's speech utilizing the speech recognition system during the first remote session, **modifying the speech recognition system** according to the sample thereby forming a speaker-specific modified speech recognition system, storing a representation of the speaker-specific modified speech recognition system in association with the identification of the speaker, and using the representation of the speaker-specific modified speech recognition system to recognize speech during a subsequent remote session with the speaker. As discussed above, Kanevsky teaches a system to modify a speaker identification process. Kanevsky does not teach a system to modify a speech recognition system. Further, Kanevsky does not teach a system to modify and subsequently store a speaker-specific speech recognition system. For at least these reasons, Claim 1 is allowable over the teachings of Kanevsky.

Claims 2-10 and 12-16 are each dependent upon the independent Claim 1. As discussed above, the independent Claim 1 is allowable over the teachings of Kanevsky. Accordingly, Claims 2-10 and 12-16 are each also allowable as being dependent upon an allowable base claim.

Within the Office Action, Claims 17-26, 28-40, and 42-58 have been rejected as having similar limitations as Claims 1-10 and 12-16. The Applicant respectfully traverses this rejection for at least the same reasons as discussed above pertaining to Claims 1-10 and 12-16.

For the reasons given above, Applicant respectfully submits that all of the remaining claims are in a condition for allowance, and allowance at an early date would be appreciated. Should the Examiner have any questions or comments, he is encouraged to call the undersigned at (408) 530-9700 to discuss the same so that any outstanding issues can be expeditiously resolved.

Respectfully submitted,
HAVERSTOCK & OWENS LLP

Dated: January 23, 2002

By: Jonathan O. Owens
Jonathan O. Owens
Reg. No. 37,902
Attorneys for Applicants

CERTIFICATE OF MAILING (37 CFR § 1.8(u))

I hereby certify that this paper (along with any referred to as being attached or enclosed) is being deposited with the U.S. Postal Service on the date shown below with sufficient postage as first class mail in an envelope addressed to the: Assistant Commissioner for Patents, Washington D.C. 20231

HAVERSTOCK & OWENS LLP

Date: 1-23-02 By: Sharon A. McCormick